

( i i ) MOLECULE TYPE: DNA (genomic)

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:4:

AGATCATCTC TGCCTGAGTA TCTT

24

( 2 ) INFORMATION FOR SEQ ID NO:5:

( i ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 24 bases
- ( B ) TYPE: nucleic acid
- ( C ) STRANDEDNESS: single
- ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: DNA (genomic)

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CCACCCATGG CAAATTCCAT GGCA

24

( 2 ) INFORMATION FOR SEQ ID NO:6:

( i ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 24 bases
- ( B ) TYPE: nucleic acid
- ( C ) STRANDEDNESS: single
- ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: DNA (genomic)

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:6:

TCTAGACGGC AGGTCAGGTC CACC

24

( 2 ) INFORMATION FOR SEQ ID NO:7:

( i ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 12 amino acids
- ( B ) TYPE: amino acid
- ( C ) STRANDEDNESS: single
- ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: protein

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Asp Asp Ile Asn Pro Thr Val Leu Leu Lys Glu Arg  
1 5 10

( 2 ) INFORMATION FOR SEQ ID NO:8:

( i ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 23 bases
- ( B ) TYPE: nucleic acid
- ( C ) STRANDEDNESS: single
- ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: DNA (genomic)

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:8:

CTGCGATGCT CGCCCGCGCC CTG

23

( 2 ) INFORMATION FOR SEQ ID NO:9:

( i ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 24 bases
- ( B ) TYPE: nucleic acid
- ( C ) STRANDEDNESS: single
- ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: DNA (genomic)

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CTTCTACAGT TCAGTCGAAC GTTC

24

## ( 2 ) INFORMATION FOR SEQ ID NO:10:

## ( i ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 604 amino acids
- ( B ) TYPE: amino acid
- ( C ) STRANDEDNESS: single
- ( D ) TOPOLOGY: linear

## ( ii ) MOLECULE TYPE: protein

## ( xi ) SEQUENCE DESCRIPTION: SEQ ID NO:10:

```

Met Leu Ala Arg Ala Leu Leu Leu Cys Ala Val Leu Ala Leu Ser His
1      5      10      15
Thr Ala Asn Pro Cys Cys Ser His Pro Cys Gln Asn Arg Gly Val Cys
20      25      30
Met Ser Val Gly Phe Asp Gln Tyr Lys Cys Asp Cys Thr Arg Thr Gly
35      40      45
Phe Tyr Gly Glu Asn Cys Ser Thr Pro Glu Phe Leu Thr Arg Ile Lys
50      55      60
Leu Phe Leu Lys Pro Thr Pro Asn Thr Val His Tyr Ile Leu Thr His
65      70      75      80
Phe Lys Gly Phe Trp Asn Val Val Asn Asn Ile Pro Phe Leu Arg Asn
85      90      95
Ala Ile Met Ser Tyr Val Leu Thr Ser Arg Ser His Leu Ile Asp Ser
100     105     110
Pro Pro Thr Tyr Asn Ala Asp Tyr Gly Tyr Lys Ser Trp Glu Ala Phe
115     120     125
Ser Asn Leu Ser Tyr Tyr Thr Arg Ala Leu Pro Pro Val Pro Asp Asp
130     135     140
Cys Pro Thr Pro Leu Gly Val Lys Gly Lys Lys Gln Leu Pro Asp Ser
145     150     155     160
Asn Glu Ile Val Glu Lys Leu Leu Leu Arg Arg Lys Phe Ile Pro Asp
165     170
Pro Gln Gly Ser Asn Met Met Phe Ala Phe Phe Ala Gln His Phe Thr
180     185     190
His Gln Phe Phe Lys Thr Asp His Lys Arg Gly Pro Ala Phe Thr Asn
195     200     205
Gly Leu Gly His Gly Val Asp Leu Asn His Ile Tyr Gly Glu Thr Leu
210     215     220
Ala Arg Gln Arg Lys Leu Arg Leu Phe Lys Asp Gly Lys Met Lys Tyr
225     230     235     240
Gln Ile Ile Asp Gly Glu Met Tyr Pro Pro Thr Val Lys Asp Thr Gln
245     250     255
Ala Glu Met Ile Tyr Pro Pro Gln Val Pro Glu His Leu Arg Phe Ala
260     265     270
Val Gly Gln Glu Val Phe Gly Leu Val Pro Gly Leu Met Met Tyr Ala
275     280     285
Thr Ile Trp Leu Arg Glu His Asn Arg Val Cys Asp Val Leu Lys Gln
290     295     300
Glu His Pro Glu Trp Gly Asp Glu Gln Leu Phe Gln Thr Ser Arg Leu
305     310     315     320
Ile Leu Ile Gly Glu Thr Ile Lys Ile Val Ile Glu Asp Tyr Val Gln
325     330     335
His Leu Ser Gly Tyr His Phe Lys Leu Lys Phe Asp Pro Glu Leu Leu

```

-continued

340										345										350									
Phe	Asn	Lys	Gln	Phe	Gln	Tyr	Gln	Asn	Arg	Ile	Ala	Ala	Glu	Phe	Asn														
		355					360					365																	
Thr	Leu	Tyr	His	Trp	His	Pro	Leu	Leu	Pro	Asp	Thr	Phe	Gln	Ile	His														
	370					375					380																		
Asp	Gln	Lys	Tyr	Asn	Tyr	Gln	Gln	Phe	Ile	Tyr	Asn	Asn	Ser	Ile	Leu														
	385				390					395					400														
Leu	Glu	His	Gly	Ile	Thr	Gln	Phe	Val	Glu	Ser	Phe	Thr	Arg	Gln	Ile														
				405					410					415															
Ala	Gly	Arg	Val	Ala	Gly	Gly	Arg	Asn	Val	Pro	Pro	Ala	Val	Gln	Lys														
			420					425					430																
Val	Ser	Gln	Ala	Ser	Ile	Asp	Gln	Ser	Arg	Gln	Met	Lys	Tyr	Gln	Ser														
		435				440						445																	
Phe	Asn	Glu	Tyr	Arg	Lys	Arg	Phe	Met	Leu	Lys	Pro	Tyr	Glu	Ser	Phe														
	450					455					460																		
Glu	Glu	Leu	Thr	Gly	Glu	Lys	Glu	Met	Ser	Ala	Glu	Leu	Glu	Ala	Leu														
	465				470					475																			
Tyr	Gly	Asp	Ile	Asp	Ala	Val	Glu	Leu	Tyr	Pro	Ala	Leu	Leu	Val	Glu														
				485					490					495															
Lys	Pro	Arg	Pro	Asp	Ala	Ile	Phe	Gly	Glu	Thr	Met	Val	Glu	Val	Gly														
			500					505					510																
Ala	Pro	Phe	Ser	Leu	Lys	Gly	Leu	Met	Gly	Asn	Val	Ile	Cys	Ser	Pro														
			515				520					525																	
Ala	Tyr	Trp	Lys	Pro	Ser	Thr	Phe	Gly	Gly	Glu	Val	Gly	Phe	Gln	Ile														
	530					535					540																		
Ile	Asn	Thr	Ala	Ser	Ile	Gln	Ser	Leu	Ile	Cys	Asn	Asn	Val	Lys	Gly														
	545				550					555					560														
Cys	Pro	Phe	Thr	Ser	Phe	Ser	Val	Pro	Asp	Pro	Glu	Leu	Ile	Lys	Thr														
				565					570					575															
Val	Thr	Ile	Asn	Ala	Ser	Ser	Ser	Arg	Ser	Gly	Leu	Asp	Asp	Ile	Asn														
			580					585					590																
Pro	Thr	Val	Leu	Leu	Lys	Glu	Arg	Ser	Thr	Glu	Leu																		
		595					600																						

## ( 2 ) INFORMATION FOR SEQ ID NO:11:

## ( i ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 3387 bases
- ( B ) TYPE: nucleic acid
- ( C ) STRANDEDNESS: single
- ( D ) TOPOLOGY: linear

## ( ii ) MOLECULE TYPE: DNA (genomic)

## ( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:11:

```

GTCCAGGAAC TCCTCAGCAG CGCCTCCTTC AGTCCACAG CCAGACGCCC TCAGACAGCA      60
AAGCCTACCC CCGCGCCGCG CCCTGCCCCG CGCTGCGATG CTCGCCCCGCG CCCTGCTGCT      120
GTGCGCGGTC CTGGCGCTCA GCCATACAGC AAATCCTTGC TGTTCACCAC CATGTCAAAA      180
CCGAGGTGTA TGTATGAGTG TGGGATTTGA CCAGTATAAG TCGATTGTA CCCGGACAGG      240
ATTCTATGGA GAAAAC TGCT CAACACCGGA ATTTTGTACA AGAATAAAAT TATTTCTGAA      300
ACCCACTCCA AACACAGTGC ACTACATACT TACCCACTTC AAGGGATTTT GGAACGTTGT      360
GAATAACATT CCCTTCCTTC GAAATGCAAT TATGAGTTAT GTGTTGACAT CCAGATCACA      420
TTTGATTGAC AGTCCACCAA CTTACAATGC TGA CTATGGC TACAAAAGCT GGAAGCCTT      480
CTCTAACCTC TCCTATTATA CTAGAGCCCT TCCTCCTGTG CCTGATGATT GCCCGACTCC      540

```

CTTGGGTGTC	AAAGGTAAAA	AGCAGCTTCC	TGATTCAAAT	GAGATTGTGG	AAAAATTGCT	600
TCTAAGAAGA	AAGTTCATCC	CTGATCCCCA	GGGCTCAAAC	ATGATGTTTG	CATTCTTTGC	660
CCAGCACTTC	ACGCACCAGT	TTTTCAAGAC	AGATCATAAG	CGAGGGCCAG	CTTTCACCAA	720
CGGGCTGGGC	CATGGGGTGG	ACTTAAATCA	TATTTACGGT	GAAACTCTGG	CTAGACAGCG	780
TAAACTGCGC	CTTTTCAAGG	ATGGAAAAAT	GAAATATCAG	ATAATTGATG	GAGAGATGTA	840
TCCTCCCACA	GTCAAAGATA	CTCAGGCAGA	GATGATCTAC	CCTCCTCAAG	TCCCTGAGCA	900
TCTACGGTTT	GCTGTGGGGC	AGGAGGTCTT	TGOTCTGGTG	CCTGGTCTGA	TGATGTATGC	960
CACAATCTGG	CTGCGGGAAC	ACAACAGAGT	ATGTGATGTG	CTTAAACAGG	AGCATCCTGA	1020
ATGGGGTGAT	GAGCAGTTGT	TCCAGACAAG	CAGGCTAATA	CTGATAGGAG	AGACTATTAA	1080
GATTGTGATT	GAAGATTATG	TGCAACACTT	GAGTGGCTAT	CACTTCAAAC	TGAAATTTGA	1140
CCCAGAACTA	CTTTTCAACA	AACAATTCCA	GTACCAAAAT	CGTATTGCTG	CTGAATTTAA	1200
CACCTCTAT	CACCTGGCATC	CCCTTCTGCC	TGACACCTTT	CAAATTCATG	ACCAGAAATA	1260
CAACTATCAA	CAGTTTATCT	ACAACAACCTC	TATATTGCTG	GAACATGGAA	TTACCCAGTT	1320
TGTTGAATCA	TTCACCAGGC	AAATTGCTGG	CAGGGTTGCT	GGTGGTAGGA	ATGTTCCACC	1380
CGCAGTACAG	AAAGTATCAC	AGGCTTCCAT	TGACCAGAGC	AGGCAGATGA	AATACCAGTC	1440
TTTTAATGAG	TACCGCAAAAC	GCTTTATGCT	GAAGCCCTAT	GAATCATTTG	AAGAACTTAC	1500
AGGAGAAAAAG	GAAATGTCTG	CAGAGTTGGA	AGCACTCTAT	GGTGACATCG	ATGCTGTGGA	1560
GCTGTATCCT	GCCCTTCTGG	TAGAAAAGCC	TCGGCCAGAT	GCCATCTTTG	GTGAAACCAT	1620
GGTAGAAGTT	GGAGCACCAT	TCTCCTTGAA	AGGACTTATG	GGTAATGTTA	TATGTTCTCC	1680
TGCCTACTGG	AAGCCAAGCA	CTTTTGGTGG	AGAAGTGGGT	TTTCAAATCA	TCAACACTGC	1740
CTCAATTCAG	TCTCTCATCT	GCAATAACGT	GAAGGGCTGT	CCCTTTACTT	CATTCAGTGT	1800
<del>TCCAGATCCA GAGCTCATTA AAACAGTCAC CATCAATGCA AGTTCTTCCC GCTCCGGACT</del>						<del>1860</del>
AGATGATATC	AATCCCACAG	TACTACTAAA	AGAACGGTCG	ACTGAACTGT	AGAAGTCTAA	1920
TGATCATATT	TATTTATTTA	TATGAACCAT	GTCTATTAAT	TTAATTATTT	AATAATATTT	1980
ATATTAAACT	CCTTATGTTA	CTTAACATCT	TCTGTAACAG	AAGTCAGTAC	TCCTGTTGCG	2040
GAGAAAGGAG	TCATACTTGT	GAAGACTTTT	ATGTCACTAC	TCTAAAGATT	TTGCTGTTGC	2100
TGTTAAOITT	GGAAAACAOT	TTTTATTCTG	TTTTATAAAC	CAGAGAGAAA	TGAGTTTTGA	2160
CGTCTTTTTA	CTTGAATTTT	AACCTTATAT	ATAAGGACGA	AAGTAAAGAT	GTTTGAATAC	2220
TTAAACACTA	TCACAAGATG	CCAAAATGCT	GAAAGTTTTT	ACACTGTCTG	TGTTTCCAAT	2280
GCATCTTCCA	TGATGCATTA	GAAGTAACTA	ATGTTTGAAA	TTTTAAAGTA	CTTTTGGGTA	2340
TTTTTCTGTC	ATCAAACAAA	ACAAGGTATCA	GTGCATTATT	AAATGAATAT	TTAAATTAGA	2400
CATTACCAGT	AATTTTCATGT	CTACTTTTTA	AAATCAGCAA	TGAAACAATA	ATTTGAAATT	2460
TCTAAATTCA	TAGGGTAGAA	TCACCTGTAA	AAGCTTGTTT	GATTTCTTAA	AGTTATTAAA	2520
CTTGACATA	TACCAAAAAG	AAGCTGTCTT	GGATTTAAAT	CTGTAAAATC	AGATGAAATT	2580
TTACTACAAT	TGCTTGTTAA	AATATTTTAT	AAGTGATGTT	CCTTTTTTCAC	CAAGAGTATA	2640
AACCTTTTTA	GTGTGACTGT	TAAAACCTCC	TTTTAAATCA	AAATGCCAAA	TTTATTAAGG	2700
TGGTGGAGCC	ACTGCAGTGT	TATCTCAAAA	TAAGAATATC	CTGTTGAGAT	ATTCCAGAAT	2760
CTGTTTATAT	GGCTGGTAAC	ATGTAAAAAC	CCCATAACCC	CGCCAAAAGG	GGTCTACCC	2820
TTGAACATAA	AGCAATAACC	AAAGGAGAAA	AGCCCAAAAT	ATTGGTTCCA	AATTTAGGGT	2880
TTAAACTTTT	TGAAGCAAAC	TTTTTTTTTAG	CCTTGTGCAC	TGCAGACCTG	GTA CT CAGAT	2940

```

TTTGCTATGA GGTAAATGAA GTACCAAGCT GTGCTTGAAT AACGATATGT TTTCTCAGAT 3000
TTTCTGTTGT ACAGTTTAAAT TTAGCAGTCC ATATCACATT GCAAAAGTAG CAATGACCTC 3060
ATAAAATACC TCTTCAAAAT GCTTAAATTC ATTTACACACA TTAATTTTAT CTCAGTCTTG 3120
AAGCCAATTC AGTAGGTGCA TTGGAATCAA GCCTGGCTAC CTGCATGCTG TTCCTTTTCT 3180
TTTCTTCTTT TAGCCATTTT GCTAAGAGAC ACAGTCTTCT CAAACACTTC GTTTCTCCTA 3240
TTTTGTTTTA CTAGTTTTAA GATCAGAGTT CACTTTCTTT GGACTCTGCC TATATTTTCT 3300
TACCTGAACT TTTGCAAGTT TTCAGGTAAA CCTCAGCTCA GGACTGCTAT TTAGCTCCTC 3360
TTAAGAAAGAT TAAAAAAAAA AAAAAAAG 3387

```

## ( 2 ) INFORMATION FOR SEQ ID NO:12:

- ( i ) SEQUENCE CHARACTERISTICS:
- ( A ) LENGTH: 21 bases
  - ( B ) TYPE: nucleic acid
  - ( C ) STRANDEDNESS: single
  - ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: DNA (genomic)

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:12:

```

CCTTCCTTCG AAATGCAATT A 21

```

## ( 2 ) INFORMATION FOR SEQ ID NO:13:

- ( i ) SEQUENCE CHARACTERISTICS:
- ( A ) LENGTH: 21 bases
  - ( B ) TYPE: nucleic acid
  - ( C ) STRANDEDNESS: single
  - ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: DNA (genomic)

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:13:

```

AAACTGATGC GTGAAGTGCT G 21

```

## ( 2 ) INFORMATION FOR SEQ ID NO:14:

- ( i ) SEQUENCE CHARACTERISTICS:
- ( A ) LENGTH: 21 bases
  - ( B ) TYPE: nucleic acid
  - ( C ) STRANDEDNESS: single
  - ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: DNA (genomic)

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:14:

```

GAGATTGTGG GAAAATTGCT T 21

```